RECOMBINANT C140 RECEPTOR, ITS AGONISTS AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR

Technical Field

The invention relates to a newly discovered receptor which is a member of the G-protein-coupled receptor superfamily. The receptor is expressed in endothelial cells in blood vessels. Avoidance of effects on this receptor is an essential element in limiting side effects of drugs which are administered to stimulate other receptors in this family. The invention also relates to nucleic acid sequences encoding the receptor protein or peptide.

Background Art

Responses of animals to many therapeutic and prophylactic drugs are mediated through receptors which reside on cell surfaces. One class of such receptors comprises the G-protein-coupled receptors, whose physiological effect is mediated by a three-subunit protein complex, called G-proteins, that binds to this type of receptor with the subsequent release of a subunit, thus setting in motion additional intracellular events. Receptors of this subclass include, among others, adrenergic receptors, neuropeptide receptors, the thrombin receptor and the C140 receptor which is the subject of the herein invention. This class of receptor is characterized by the presence of seven transmembrane regions which anchor the receptor within the cell surface.

It is the elusive goal of the designers of therapeutic substances to effect a desired response in a

subject in the absence of side effects. Accordingly, pharmaceuticals designed to target a specific receptor, such as the thrombin receptor, should react with the thrombin receptor specifically and have no effect on related receptors. The C140 receptor of the present invention may be involved in controlling vascular pressure, and inadvertent stimulation or blocking of this receptor would have unpredictable and therefore undesirable results. It is therefore useful to determine in advance whether therapeutic reagents designed to target, for example, the thrombin receptor will or will not have the undesired side effect of reactivity with the C140 receptor. By providing the recombinant materials for the production of the C140 receptor in convenient assay systems, as well as agonist and antagonist reagents for use in this assay, the invention makes possible the prior determination of the presence or absence of the side effect of reactivity with the C140 receptor in candidate pharmaceuticals. This side effect will usually be undesired as it is believed that the C140 receptor responds to enzymes such as serine proteases associated with trauma and immune disturbances.

Disclosure of the Invention

The invention provides methods and materials useful in assay systems to determine the propensity of candidate pharmaceuticals to exert undesirable side effects. The isolation, recombinant production and characterization of the C140 receptor permits the design of assay systems using the receptor as a substrate and using agonists and

antagonists for the receptor as control reagents in the assay.

Thus, in one aspect, the invention is directed to recombinant materials associated with the production of C140 receptor. These include, for example, transfected cells which can be cultured so as to display the C140 receptor on their surfaces, and thus provide an assay system for the interaction of materials with the native C140 receptor. In general, the limitations on the host cells useful in these assay systems are that the cells have the appropriate mechanism to display the receptor on their surfaces and contain the G-protein as mediator to the intracellular response. (However assays which merely assess binding do not require the G-protein.) Most animal cells meet these requirements.

In another aspect, the invention is directed to C140 receptor agonists which mimic the activated form of the extracellular portion of the receptor protein. These agonists are useful as control reagents in the above-mentioned assays to verify the workability of the assay system. In addition, agonists for the C140 receptor may exhibit hypotensive effects *in vivo*. Accordingly, the agonists may be also, themselves, useful as antihypertensives.

In still another aspect, the invention is directed to C140 receptor antagonists. These antagonists comprise modified forms of the C140 receptor agonist peptides that lack the essential features required for activation of the receptor. These antagonists bind to receptor, do not activate it, and prevent receptor activation by agonists and the native receptor-binding ligand.

A second group of antagonists includes antibodies designed to bind specific portions of the receptor protein. In general, these are monoclonal antibody preparations which are highly specific for any desired region of the C140 receptor. The antibodies of the invention are also useful in immunoassays for the receptor protein, for example, in assessing successful expression of the gene in recombinant systems.

Another aspect of the invention is to provide nucleic acids encoding such a C140 receptor polypeptide and to use this nucleic acid to produce the polypeptide in recombinant cell culture for diagnostic use or for potential therapeutic use in hemostatic or immune response regulation.

In still other aspects, the invention provides an isolated nucleic acid molecule encoding a C140 receptor, labeled or unlabeled, and a nucleic acid sequence that is complementary to, or hybridizes under stringent conditions to, a nucleic acid sequence encoding a C140 receptor. The isolated nucleic acid molecule of the present invention excludes nucleic acid sequences which encode, or are complementary to nucleic acid sequences encoding, other known G protein-coupled receptors which are not C140 receptors, such as adrenergic receptors, neuropeptide receptors, thrombin receptors, and the like.

In addition, the invention provides a replicable vector comprising a nucleic acid molecule encoding a C140 receptor operably linked to control sequences recognized by a host transformed by the vector; host cells transformed with the vector; and a method of using a

nucleic acid molecule encoding a C140 receptor to effect the production of a C140 receptor, comprising expressing the nucleic acid molecule in a culture of the transformed host cells and recovering a C140 receptor from the host cell culture. The nucleic acid sequence is also useful in hybridization assays for C140 receptor-encoding nucleic acid molecules.

In still further embodiments, the invention provides a method for producing C140 receptors comprising inserting into the DNA of a cell containing the nucleic acid sequence encoding a C140 receptor a transcription modulatory element in sufficient proximity and orientation to the C140 receptor coding sequence to influence transcription thereof, with an optional further step comprising culturing the cell containing the transcription modulatory element and the C140 receptor-encoding nucleic acid sequence.

In still further embodiments, the invention provides a cell comprising a nucleic acid sequence encoding a C140 receptor and an exogenous transcription modulatory element in sufficient proximity and orientation to the above coding sequence to influence transcription thereof; and a host cell containing the nucleic acid sequence encoding a C140 receptor operably linked to exogenous control sequences recognized by the host cell.

Still further is provided a method for obtaining cells having increased or decreased transcription of the nucleic acid molecule encoding a C140 receptor, comprising:

(a) providing cells containing the nucleic acid molecule;

(b) introducing into the cells a transcription modulating element; and

(c) screening the cells for a cell in which the transcription of the nucleic acid molecule is increased or decreased.

In another aspect, the invention is related to assay systems which utilize recombinant C140 receptor to screen for agonist and antagonist activity of candidate drugs. This assay is especially useful in assuring that these therapeutic agents do not have undesired side effects caused by activation or inhibition of the C140 receptor. In some cases agonist activity at this receptor system may have therapeutic utility. Some of these assay systems include the use of the agonist peptides as positive controls. The assay can also be used to screen for antagonists which inhibit the agonistic effect.

Another aspect of the invention relates to the diagnosis of conditions characterized by activation of the C140 receptor by detection in fluids, such as blood or urine, of the peptide cleaved from the C140 receptor when the receptor is activated. Another diagnostic method included in the invention is visualization of the activated forms of receptor by localizing an imaging agent to activated receptor *in situ* using antibodies specific to the activated receptor.

Yet another aspect of this invention relates to the therapeutic, prophylactic and research uses of various techniques to block or modulate the expression of a C140 receptor by interfering with the transcription of translation of a DNA or RNA molecule encoding the C140 receptor. This includes a method to inhibit or regulate

expression of C140 receptors in a cell comprising providing to the cell an oligonucleotide molecule which is antisense to, or forms a triple helix with, C140 receptor-encoding DNA or with DNA regulating expression of C140 receptor-encoding DNA, in an amount sufficient to inhibit or regulate expression of the C140 receptors, thereby inhibiting or regulating their expression. Also included is a method to inhibit or regulate expression of C140 receptors in a subject, comprising administering to the subject an oligonucleotide molecule which is antisense to, or forms a triple helix with, C140 receptor-encoding DNA or with DNA regulating expression of C140 receptor-encoding DNA, in an amount sufficient to inhibit or regulate expression of the C140 receptors in the subject, thereby inhibiting or regulating their expression. The antisense molecule or triple helix-forming molecule in the above methods is preferably a DNA or RNA oligonucleotide.

Additional aspects of the invention are directed to pharmaceutical compositions containing the agonists and antagonists of the invention. The agonists of the invention are antihypertensives; conversely, the antagonists can elevate blood pressure if desired. Other aspects of the invention include a pharmaceutical composition useful for inhibiting or regulating C140 receptor expression in a cell or in a subject at the level of transcription or translation, which composition comprises an antisense or triple helix-forming molecule as described above which corresponds to a portion of the sequence of the C140 receptor-coding nucleic acid.

Brief Description of the Drawings

Figures 1A-1B show the DNA and deduced amino acid sequence of murine C140 receptor.

Figures 2A-2B show the DNA and deduced amino acid sequence of human C140 receptor.

Figure 3 shows a comparison of amino acid sequences for the human C140 receptor and murine C140 receptor.

Figure 4 shows a proposed model of C140 receptor activation based on the deduced amino acid sequence.

Figure 5 shows a comparison of amino acid sequences for the mouse C140 receptor and the human thrombin receptor.

Figure 6 shows the results of Northern Blot to detect the presence of mRNA encoding C140 receptor in various mouse tissues.

Figure 7 shows a trace of blood pressure demonstrating the *in vivo* hypotensive effect of a C140 agonist peptide.

Figures 8a-8c show blood vessel dilation in rat femoral vein induced by a C140 receptor agonist peptide. Figure 8a shows these results in the immobilized vein; Figure 8b shows these results for the immobilized vein depleted of endothelial cells.

Figures 9a-9c show the results of an assay for activation of the C140 receptor, expressed in frog oocytes, by plasmin, kallikrein, or trypsin. Figure 9a shows the results for plasmin; Figure 9b shows the results for kallikrein; Figure 9c shows the results for trypsin.

Figures 10A-10B show the nucleotide sequence and deduced amino acid sequence of a cDNA clone encoding murine C140 receptor.

Figures 11A-11B show the nucleotide sequence and deduced amino acid sequence of a cDNA clone encoding human C140 receptor.

Figure 12 shows the results of *in situ* hybridization of a sectioned newborn mouse with mouse C140 receptor probes.

Figure 13 shows a Northern blot of total RNA from human cell lines hybridized to a human C140 receptor probe.

Modes of Carrying Out the Invention

The characteristics of the C140 receptor elucidated by the invention herein are summarized in Figures 1A/1B-4. Figures 1A-1B shows the complete DNA sequence of the clone encoding the murine receptor, along with the deduced amino acid sequence. As used herein, the "C140 receptor" refers to receptor in any animal species corresponding to the murine receptor contained in clone C140 described in Example 1 herein. Using the native DNA encoding the murine form of this receptor, the corresponding receptors in other species, including humans, as illustrated herein, may be obtained. Figures 2A-2B shows the corresponding DNA and deduced amino acid sequence of the human receptor.

The entire amino acid sequence of the murine receptor contains 395 amino acids, including a 27 amino acid signal peptide which, when cleaved, results in a 368 amino acid mature receptor protein. Similarly, the human receptor is encoded by an open reading frame corresponding to 398 amino acids including a probable 29 amino acid signal

peptide sequence resulting in a 369 amino acid mature receptor protein, as shown in Figures 2A-2B.

Figure 3 shows a comparison of the human and murine amino acid sequences; as shown, these sequences exhibit a high degree of homology.

Hydrophobicity/hydrophilicity plots of the sequences shown in Figures 1A-1B and 2A-2B indicate that the mature C140 receptor is a member of the 7-transmembrane domain receptor family whose effect on the cell is mediated by G-protein. The mature C140 receptor has a relatively long extracellular amino acid extension containing several consensus sites for asparagine-linked glycosylation. It also contains a conserved asparagine in the first transmembrane region, the motif Leu-Ala-X-X-Asp in the second transmembrane region, a Trp in the fourth transmembrane region and a carboxy terminal tail which contains multiple serine and threonine residues. A proposed model of the *in situ* receptor is shown in Figure 4.

Referring to Figure 5, similarities to the thrombin receptor are readily seen. Figure 5 compares the amino acid sequence of murine C140 with that of thrombin receptor. It is known that the thrombin receptor is activated by proteolytic cleavage of the Arg-Ser bond at positions 41 and 42, which releases an activation peptide that permits refolding of the receptor and activation via the newly created amino terminus. In an analogous manner, the C140 receptor is activated by cleavage of the Arg-Ser bond at positions 34 and 35, also liberating an activation peptide extending from position 1 of the putative mature protein to the cleavage site. It is believed that Arg-28

is the amino terminal amino acid residue of the mature protein, so the activation peptide has the sequence RNNSKGR. This peptide could thus be used as an index for activation of C140 receptor. In any event, the precise location of the N-terminus of the mature protein is unimportant for the design of agonists or antagonists. The activation peptide is likely to be freely filtered by the kidney and possibly concentrated in the urine and can be used as an index to activation of the C140 receptor.

Release of the activation peptide permits refolding of the receptor protein to activate the receptor. This is shown schematically in Figure 4, which also shows that the conformational changes resulting from the liberation of the activation peptide and refolding results in an intracellular conformational change of the receptor. This hypothesis is confirmed by the finding that the C140 receptor can be activated by a peptide mimicking the new amino terminus created by the activation. Accordingly, mimics of the N-terminus of the new amino terminus on the activated receptor behave as agonists therefor. The importance of the first five amino acids in the newly created amino terminus in the receptor for receptor activation has also been confirmed hereinbelow.

Based on this information, and by analogy with the mechanisms underlying trypsinogen activation to trypsin and activation of the thrombin receptor, it appears that the positively charged amino group on serine that is newly exposed when the ligand cleaves the receptor plays an important role in receptor activation. Peptides based on the agonist peptide sequence that bind the C140 receptor, but which are modified to be lacking the free α-amino

group can function as antagonists of this receptor. Thus, modifications of the agonist peptides which lack the capacity for specific activating interaction serve as C140 receptor antagonists.

Ordinarily, the C140 receptors and analogs thereof claimed herein will have an amino acid sequence having at least 75% amino acid sequence identity with a "common" C140 receptor sequence (such as that disclosed in Figures 1A-1B or Figures 2A-2B), more preferably at least 80%, even more preferably at least 90%, and most preferably at least 95%. Identity or homology with respect to a common sequence is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the known C140 receptor, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity. None of N-terminal, C-terminal or internal extensions, deletions, or insertions into the C140 receptor sequence shall be construed as affecting homology.

Thus, the claimed C140 receptor and analog molecules that are the subject of this invention include molecules having the C140 receptor amino acid sequence; fragments thereof having a consecutive sequence of at least 10, 15, 20, 25, 30 or 40 amino acid residues from a common C140 receptor sequence; amino acid sequence variants of a common C140 receptor sequence wherein an amino acid residue has been inserted N- or C-terminal to, or within, the C140 receptor sequence or its fragments as defined above; amino acid sequence variants of the common C140 receptor sequence or its fragment as defined above

which have been substituted by another residue. C140 receptor polypeptides include those containing predetermined mutations by, *e.g.*, homologous recombination, site-directed or PCR mutagenesis, and C140 receptor polypeptides of other animal species, including but not limited to rabbit, rat, murine, porcine, bovine, ovine, equine and non-human primate species, and alleles or other naturally occurring variants of the C140 receptor of the foregoing species and of human sequences; derivatives of the commonly known C140 receptor or its fragments wherein the C140 receptor or its fragments have been covalently modified by substitution, chemical, enzymatic, or other appropriate means with a moiety other than a naturally occurring amino acid (for example a detectable moiety such as an enzyme or radioisotope); glycosylation variants of C140 receptor (insertion of a glycosylation site or deletion of any glycosylation site by deletion, insertion or substitution of appropriate amino acid); and soluble forms of C140.

The novel proteins and peptides of the present invention are preferably those which share a common biological activity with the C140 receptor, including but not limited to an effector or receptor function or cross-reactive antigenicity. Such fragments and variants exclude any C140 receptor polypeptide heretofore made public, including any known protein or polypeptide of any animal species, which is otherwise anticipatory under 35 U.S.C. §102 as well as polypeptides obvious over such known protein or polypeptides under 35 U.S.C. §103. Specifically, the present C140 receptor proteins, analogs, fragments and variants exclude other known G protein-

coupled receptors which are not C140 receptors, such as adrenergic receptors, neuropeptide receptors, thrombin receptors, and the like.

Compounds of the Invention

The nomenclature used to describe the peptide compounds of the invention follows the conventional practice where the N-terminal amino group is assumed to be to the left and the carboxy group to the right of each amino acid residue in the peptide. In the formulas representing selected specific embodiments of the present invention, the amino- and carboxy-terminal groups, although often not specifically shown, will be understood to be in the form they would assume at physiological pH values, unless otherwise specified. Thus, the N-terminal H^{+}_{2} and C-terminal O^{-} at physiological pH are understood to be present though not necessarily specified and shown, either in specific examples or in generic formulas. Free functional groups on the side chains of the amino acid residues can also be modified by amidation, acylation or other substitution, which can, for example, change the solubility of the compounds without affecting their activity.

In the peptides shown, each gene-encoded residue, where appropriate, is represented by a single letter designation, corresponding to the trivial name of the amino acid, in accordance with the following conventional list:

Amino Acid	One-Letter Symbol	Three-letter Symbol
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp
Cysteine	C	Cys
Glutamine	Q	Gln
Glutamic acid	E	Glu
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

The amino acids not encoded genetically are abbreviated as indicated in the discussion below.

In the specific peptides shown in the present application, the L-form of any amino acid residue having an optical isomer is intended unless the D-form is expressly indicated by a dagger superscript (†).

The compounds of the invention are peptides which are partially defined in terms of amino acid residues of designated classes. Amino acid residues can be generally subclassified into four major subclasses as follows:

Acidic: The residue has a negative charge due to loss of H ion at physiological pH and the residue is attracted by aqueous solution so as to seek the surface positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium at physiological pH.

Basic: The residue has a positive charge due to association with H ion at physiological pH and the residue is attracted by aqueous solution so as to seek the surface positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium at physiological pH.

Neutral/nonpolar: The residues are not charged at physiological pH and the residue is repelled by aqueous solution so as to seek the inner positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium. These residues are also designated "hydrophobic" herein.

Neutral/polar: The residues are not charged at physiological pH, but the residue is attracted by aqueous solution so as to seek the outer positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium.

It is understood, of course, that in a statistical collection of individual residue molecules some molecules will be charged, and some not, and there will be an attraction for or repulsion from an aqueous medium to a

greater or lesser extent. To fit the definition of "charged," a significant percentage (at least approximately 25%) of the individual molecules are charged at physiological pH. The degree of attraction or repulsion required for classification as polar or nonpolar is arbitrary and, therefore, amino acids specifically contemplated by the invention have been classified as one or the other. Most amino acids not specifically named can be classified on the basis of known behavior.

Amino acid residues can be further subclassified as cyclic or noncyclic, and aromatic or nonaromatic, self-explanatory classifications with respect to the side chain substituent groups of the residues, and as small or large. The residue is considered small if it contains a total of 4 carbon atoms or less, inclusive of the carboxyl carbon. Small residues are, of course, always nonaromatic.

For the naturally occurring protein amino acids, subclassification according to the foregoing scheme is as follows.

Acidic: Aspartic acid and Glutamic acid;
Basic/noncyclic: Arginine, Lysine;
Basic/cyclic: Histidine;
Neutral/polar/small: Glycine, serine, cysteine;
Neutral/nonpolar/small: Alanine;
Neutral/polar/large/nonaromatic: Threonine, Asparagine, Glutamine;
Neutral/polar/large aromatic: Tyrosine;
Neutral/nonpolar/large/nonaromatic: Valine, Isoleucine, Leucine, Methionine;
Neutral/nonpolar/large/aromatic: Phenylalanine, and Tryptophan

The gene-encoded secondary amino acid proline, although technically within the group neutral/nonpolar/large/ cyclic and nonaromatic, is a special case due to its known effects on the secondary conformation of peptide chains, and is not, therefore, included in this defined group.

Certain commonly encountered amino acids, which are not encoded by the genetic code, include, for example, beta-alanine (beta-Ala), or other omega-amino acids, such as 3-amino propionic, 2,3-diamino propionic (2,3-diaP), 4-amino butyric and so forth, alpha-aminisobutyric acid (Aib), sarcosine (Sar), ornithine (Orn), citrulline (Cit), t-butylalanine (t-BuA), t-butylglycine (t-BuG), N-methylisoleucine (N-MeIle), phenylglycine (Phg), and cyclohexylalanine (Cha), norleucine (Nle), cysteic acid (Cya) 2-naphthylalanine (2-Nal); 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid (Tic); β-2-thienylalanine (Thi); and methionine sulfoxide (MSO). These also fall conveniently into particular categories.

Based on the above definitions,

Sar, beta-Ala, 2,3-diaP and Aib are neutral/nonpolar/ small;

t-BuA, t-BuG, N-MeIle, Nle, Mvl and Cha are neutral/nonpolar/large/nonaromatic;

Orn is basic/noncyclic;

Cya is acidic;

Cit, Acetyl Lys, and MSO are neutral/polar/ large/nonaromatic; and

Phg, Nal, Thi and Tic are neutral/nonpolar/large/ aromatic.

The various omega-amino acids are classified according to size as neutral/nonpolar/small (beta-Ala, i.e., 3-aminopropionic, 4-aminobutyric) or large (all others).

Other amino acid substitutions of those encoded in the gene can also be included in peptide compounds within the scope of the invention and can be classified within this general scheme according to their structure.

All of the compounds of the invention, when an amino acid forms the C-terminus, may be in the form of the pharmaceutically acceptable salts or esters. Salts may be, for example, Na^+, K^+, Ca^{+2}, Mg^{+2} and the like; the esters are generally those of alcohols of 1-6C.

In all of the peptides of the invention, one or more amide linkages (-CO-NH-) may optionally be replaced with another linkage which is an isostere such as $-CH_2NH-$, $-CH_2S-$, $-CH_2CH_2$, -CH=CH- (cis and trans), $-COCH_2-$, $-CH(OH)CH_2-$ and $-CH_2SO-$. This replacement can be made by methods known in the art. The following references describe preparation of peptide analogs which include these alternative-linking moieties: Spatola, A.F., Vega Data (March 1983), Vol. 1, Issue 3, "Peptide Backbone Modifications" (general review); Spatola, A.F., in "Chemistry and Biochemistry of Amino Acids Peptides and Proteins," B. Weinstein, eds., Marcel Dekker, New York, p. 267 (1983) (general review); Morley, J.S., Trends Pharm Sci (1980) pp. 463-468 (general review); Hudson, D., et al., Int J Pept Prot Res (1979) 14:177-185 ($-CH_2NH-$, $-CH_2CH_2-$); Spatola, A.F., et al., Life Sci (1986) 38:1243-1249 ($-CH_2-S$); Hann, M.M., J Chem Soc Perkin Trans I

(1982) 307-314 (-CH-CH-, cis and trans); Almquist, R.G., et al., J Med Chem (1980) 23:1392-1398 ($-COCH_2-$); Jennings-White, C., et al., Tetrahedron Lett (1982) 23:2533 ($-COCH_2-$); Szelke, M., et al., European Application EP 45665 (1982) CA:97:39405 (1982) ($-CH(OH)CH_2-$); Holladay, M.W., et al., Tetrahedron Lett (1983) 24:4401-4404 ($-C(OH)CH_2-$); and Hruby, V.J., Life Sci (1982) 31:189-199 ($-CH_2-S-$).

A. Agonists

The agonists of the invention comprise a series of peptides of the formula

$$AA_1-AA_2-AA_3-AA_4-AA_5-AA_6-AA_7-Z \qquad (1)$$

wherein AA_1 is a small amino acid or threonine;

AA_2 and AA_3 are each independently neutral/nonpolar/large/nonaromatic amino acids;

AA_4 is a small amino acid;

AA_5 is a basic amino acid;

AA_6 may be present or absent and, if present, is a neutral/nonpolar/large/nonaromatic amino acid;

AA_7 is absent if AA_6 is absent and may be present or absent if AA_6 is present, and is an acidic amino acid; and

Z is a substituent that does not interfere with agonist activity.

The peptide of formula 1 can be extended (shown as included in Z) at the C-terminus (but not the N-terminus)

by further amino acid sequence to comprise a noninterfering substituent.

At the C-terminus of the compounds of formula 1, the carboxyl group may be in the underivatized form or may be amidated or may be an ester; in the underivatized form the carboxyl may be as a free acid or a salt, preferably a pharmaceutically acceptable salt.

If the C-terminus is amidated, the nitrogen atom of the amido group, covalently bound to the carbonyl carbon at the C-terminus, will be NR'R', wherein each R' is independently hydrogen or is a straight or branched chain alkyl of 1-6C, such alkyls are 1-6C straight- or branched-chain saturated hydrocarbyl residues, such as methyl, ethyl, isopentyl, n-hexyl, and the like. Representatives of such amido groups are: $-NH_2$, $-NHCH_3$, $-N(CH_3)_2$, $-NHCH_2CH_3$, $-NHCH_2CH(CH_3)_2$, and $-NHCH_2CH(CH_3)CH_2CH_3$, among others. Furthermore, either or both R' may in turn optionally be substituted by one or more substituents such as, for example, -OR', -NR'R', halo, -NR'CNR'NR'R' and the like, wherein each R' is as independently defined above. Thus, Z may be -OH, or an ester (OR') or salt forms thereof, or -NR'R' wherein R' is as above defined.

Preferred embodiments of AA_1 are Ser on 2,3-diaminopropionyl (2,3-diaP). Preferred embodiments of AA_2 and AA_3 are Val, Ile, Cha and Leu. Preferred embodiments for the residues in the remainder of the compound of formula (1) are those wherein AA_4 is Gly, AA_5 is Lys, Arg or Har, AA_6, if present, is Val, Ile, Cha or Leu, and AA_7, if present, is Asp or Glu. Particularly preferred are compounds of formula (1) which are selected from the group

consisting of SLIGRLETQPPIT, SLIGRLETQPPI, SLIGRLETQPP, SLIGRLETQP, SLIGRLETQ, SLIGRLET, SLIGRLE, SLIGRL, SLIGR, SLLGKVDGTSHVT, SLLGKVDGTSHV, SLLGKVDGTSH, SLLGKVDGTS, SLLGKVDGT, SLLGKVDG, SLLGKVD, SLLGKV, SLLGK, S(Cha)IGR, S(Cha)LGK, (2,3-diaP)-IGR, (2,3-diaP)LLGK, SLLGKR-NH_2, SLIGRR-NH_2, S(Cha)LGKK-NH_2, S(Cha)IGRK-NH_2, (2,3-diaP)-LIGRK-NH_2, (2,3-diaP)-LLGKK-NH_2 and the amidated forms thereof.

B. Antagonists

Compounds of the invention which interfere with activities mediated by the C140 receptor include modified agonist peptides lacking the N-terminal serine residue; and antibodies which are immunoreactive with various critical positions on the C140 receptor.

Peptide Antagonists

The antagonists of the first group--modified agonists--can be represented by the formula:

$$X\text{-}AA_2\text{-}AA_3\text{-}AA_4\text{-}AA_5\text{-}AA_6\text{-}AA_7\text{-}Z$$

wherein X is an amino acid residue other than ser, ala, thr, cys, 2,3-diaP or gly or is a desamino or alkylated or acylated amino acid,

wherein AA_2 and AA_3 are each independently neutral/nonpolar/large/nonaromatic amino acids;

AA_4 is a small amino acid;

AA_5 is a basic amino acid;

AA_6 may be present or absent and, if present, is a neutral/nonpolar/large/nonaromatic amino acid;

AA_7 is absent if AA_6 is absent and may be present or absent if AA_6 is present, and is an acidic amino acid; and

Z is a substituent that does not interfere with agonist activity.

Preferred acyl groups are of the formula RCO- wherein R represents a straight or branched chain alkyl of 1-6C. Acetyl is particularly preferred.

Preferred embodiments of X include residues of 3-mercaptopropionic acid (Mpr), 3-mercaptovaleric acid (Mvl), 2-mercaptobenzoic acid (Mba) and S-methyl-3-mercaptopropionic acid (SMeMpr). Preferred embodiments for AA_2 through AA_7 are as described for the agonists above; Z is also as thus described.

Particularly preferred among the antagonist peptides of this class are those selected from the group consisting of Mpr-LLGK, Mpr-LIGR, Mpr-(Cha)LKG, Mpr-(Cha)IGR, Mpr-LLGKK-NH_2, Mpr-LIGRK-NH_2, Mpr-LIGRKETQP-NH_2, Mpr-LLGKKDGTS-NH_2, (n-pentyl)$_2$-N-Leu-Ile-Gly-Arg-Lys-NH_2 and (Me-N-(n-pentyl)-Leu-Ile-Gly-Arg-Lys-NH_2.

Antibodies

Antagonists which are antibodies immunoreactive with critical positions of the C140 receptor are obtained by immunization of suitable mammalian subjects with peptides containing as antigenic regions those portions of the C140 receptor intended to be targeted by the antibodies. Critical regions include the region of proteolytic cleavage, the segment of the extracellular segment

critical for activation (this includes the cleavage site), and the portions of the sequence which form the extracellular loops, in particular, that region which interacts with the N-terminus of the activated receptor extracellular region. The agonist peptides of the invention may be used as immunogens in this case.

Thus, peptides which contain the proteolytic region, namely, for example, SKGRSLIGRLET, the extracellular loops, such as those including ISY HLHGNNWVYGEALC; QTIYIPALNITTCHDVLPEEVLVGDMFNYFL; and HYFLIKTQRQSHVYA. The agonist peptides described below are also useful as immunogens.

The antibodies are prepared by immunizing suitable mammalian hosts in appropriate immunization protocols using the peptide haptens alone, if they are of sufficient length, or, if desired, or if required to enhance immunogenicity, conjugated to suitable carriers. Methods for preparing immunogenic conjugates with carriers such as BSA, KLH, or other carrier proteins are well known in the art. In some circumstances, direct conjugation using, for example, carbodiimide reagents may be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co., Rockford, IL, may be desirable to provide accessibility to the hapten. The hapten peptides can be extended at the amino or carboxy terminus with a Cys residue or interspersed with cysteine residues, for example, to facilitate linking to carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the

immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

While the polyclonal antisera produced in this way may be satisfactory for some applications, for pharmaceutical compositions, use of monoclonal preparations is preferred. Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using the standard method of Kohler and Milstein or modifications which effect immortalization of lymphocytes or spleen cells, as is generally known. The immortalized cell lines secreting the desired antibodies are screened by immunoassay in which the antigen is the peptide hapten or is the C140 receptor itself displayed on a recombinant host cell. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either *in vitro* or by production in ascites fluid.

The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonals or the polyclonal antisera which contain the immunologically significant portion can be used as antagonists, as well as the intact antibodies. Use of immunologically reactive fragments, such as the Fab, Fab', of $F(ab')_2$ fragments is often preferable, especially in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin.

The antibodies or fragments may also be produced, using current technology, by recombinant means. Regions that bind specifically to the desired regions of receptor

can also be produced in the context of chimeras with multiple species origin.

The antibodies thus produced are useful not only as potential antagonists for the receptor, filling the role of antagonist in the assays of the invention, but are also useful in immunoassays for detecting the activated receptor. As such these antibodies can be coupled to imaging agents for administration to a subject to allow detection of localized antibody to ascertain the position of C140 receptors in either activated or unactivated form. In addition, these reagents are useful _in vitro_ to detect, for example, the successful production of the C140 receptor deployed at the surface of the recombinant host cells.

Preparation of Peptide Agonists and Antagonists

The peptide agonists and antagonists of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

Preparation of C140 Receptor Nucleic Acids

C140 receptor "nucleic acid" is defined as RNA or DNA that encodes a C140 receptor, or is complementary to nucleic acid sequence encoding a C140 receptor, or hybridizes to such nucleic acid and remains stably bound to it under stringent conditions, or encodes a polypeptide sharing at least 75% sequence identity, preferably at least 80%, and more preferably at least 85%, with the translated amino acid sequences shown in Figures 3, 10A-10B or 11A-11B. It is typically at least about 10 nucleotides in length and preferably has C140 receptor biological or immunological activity, including the nucleic acid encoding an activation peptide fragment having the nucleotide sequence shown in Figure 4. Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as nucleic acids based on alternative backbone or including alternative bases whether derived from natural sources or synthesized. Such hybridizing or complementary nucleic acid, however, is defined further as being novel and unobvious over any prior art nucleic acid including that which encodes, hybridizes under stringent conditions, or is complementary to nucleic acid encoding a known G protein-coupled receptor.

"Stringent conditions" are those that (1) employ low ionic strength and high temperature for washing, for example, 0,015M NaCl/0.0015M sodium titrate/0.1% NaDodSO4 at 50o C., or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH

6.5 with 750 mM NaCl, 75 mM sodium citrate at 42o C. Another example is use of 50% formamide, 5 x SSC (0.75M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 mu g/ml), 0.1% SDS, and 10% dextran sulfate at 42o C., with washes at 42o C. in 0.2 x SSC and 0.1% SDS.

"Isolated" nucleic acid will be nucleic acid that is identified and separated from contaminant nucleic acid encoding other polypeptides from the source of nucleic acid. The nucleic acid may be labeled for diagnostic and probe purposes, using any label known and described in the art as useful in connection with diagnostic assays.

Of particular interest is a C140 receptor nucleic acid that encodes a full-length molecule, including but not necessarily the native signal sequence thereof. Nucleic acid encoding full-length protein is obtained by screening selected cDNA (not kidney) or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures to secure DNA that is complete at its 5' coding end. Such a clone is readily identified by the presence of a start codon in reading frame with the original sequence.

DNA encoding an amino acid sequence variant of a C140 receptor is prepared as described below or by a variety of methods known in the art. These methods include, but are not limited to, isolation from a natural source (in the case of naturally occurring amino acid sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette

mutagenesis of an earlier prepared variant or a non-variant version of a C140 receptor.

Techniques for isolating and manipulating nucleic acids are disclosed for example by the following documents: U.S. 5,030,576, U.S. 5,030,576 and International Patent Publications WO94/11504 and WO93/03162. See, also, Sambrook, J. *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989, and Ausubel, F.M. *et al. Current Protocols in Molecular Biology*, Vol. 2, Wiley-Interscience, New York, 1987. Disclosures of these documents are expressly incorporated herein by reference in their entireties.

Recombinant Production of C140 Receptor for Use in Assays

The invention provides recombinant materials for the production of C140 receptor for display on the surface of recombinant cells. Production of the receptor using these recombinant methods provides a useful reagent to determine the ability of a candidate drug to bind to, to activate, or to antagonize the C140 receptor. Determination of these properties is essential in evaluating the specificity of drugs intended for binding other related receptors.

For this recombinant production, a DNA sequence encoding the C140 receptor, such as those set forth in Figures 1A-1B and 2A-2B, or their substantial equivalents or their degenerate analogs, is prepared either by retrieval of the native sequence, as set forth below, or by using substantial portions of the known native sequence as probe, or can be synthesized de novo using standard

procedures. The DNA is ligated into expression vectors suitable for the desired host and transformed into compatible cells. The cells are cultured under conditions which favor the expression of the C140 receptor encoding gene and the cells displaying the receptor on the surface are harvested for use in the assays.

The host cells are typically animal cells, most typically mammalian cells. In order to be useful in the assays, the cells must have intracellular mechanisms which permit the receptor to be displayed on the cell surface in the configuration shown generally in **Figure 4** herein. If the assay uses cellular response to activated receptor as a detection system, the cells must also contain a G-protein linked mechanism for response to activation of the receptors. Most mammalian and other animal cells fulfill these qualifications.

Particularly useful cells for use in the method of the invention are *Xenopus laevis* frog oocytes, which typically utilize cRNA rather than standard recombinant expression systems proceeding from the DNA encoding the desired protein. Capped RNA (at the 5' end) is typically produced from linearized vectors containing DNA sequences encoding the receptor. The reaction is conducted using RNA polymerase and standard reagents. cRNA is recovered, typically using phenol/chloroform precipitation with ethanol and injected into the oocytes.

The animal host cells expressing the DNA encoding the C140 receptor or the cRNA-injected oocytes are then cultured to effect the expression of the encoding nucleic acids so as to produce the C140 receptor displayed in a manner analogous to that shown in **Figure 4** on their

surfaces. These cells then are used directly in assays for assessment of a candidate drug to bind, antagonize, or activate the receptor.

Assays

In one type of easily conducted assay, competition of the candidate drug for binding to the receptor with either agonist or known binding antagonist can be tested. In one method, the competing agonist or antagonist may be labeled; the labeled substance known to bind the receptor can, of course, be a synthetic peptide. In one typical protocol, varying concentrations of the candidate are supplied along with a constant concentration of labeled agonist or antagonist and the inhibition of a binding of label to the receptor can be evaluated using known techniques.

In a somewhat more sophisticated approach, the effect of candidate compounds on agonist-induced responses can be measured in the cells recombinantly expressing the C140 receptor as described below. Assay systems for the effect of activation of receptor on these cells include calcium mobilization and voltage clamp which are described herein in further detail. These assays permit an assessment of the effect of the candidate drug on the receptor activity rather than simply ability to bind to the receptor.

Agonist-induced increases in ^{45}Ca release by oocytes expressing cRNA encoding C140 receptor or other recombinant cells producing C140 receptor are assessed by published techniques (Williams, J.A., et al., Proc Natl Acad Sci USA (1988) 85:4939-4943). Briefly, intracellular calcium pools are labeled by incubating groups of 30

oocytes in 300 μl calcium-free modified Barth's solution (MBSH) containing 50 μCi $^{45}CaCl_2$ (10-40 mCi/mg Ca; Amersham) for 4 hours at RT. The labeled oocytes or cells are washed, then incubated in MBSH II without antibiotics for 90 minutes. Groups of 5 oocytes are selected and placed in individual wells in a 24-well tissue culture plate (Falcon 3047) containing 0.5 ml/well MBSH II without antibiotics. This medium is removed and replaced with fresh medium every 10 minutes; the harvested medium is analyzed by scintillation counting to determine ^{45}Ca released by the oocytes during each 10-minute incubation. The 10-minute incubations are continued until a stable baseline of ^{45}Ca release per unit time is achieved. Two additional 10-minute collections are obtained, then test medium including agonist is added and agonist-induced ^{45}Ca release determined.

Using the above assay, the ability of a candidate drug to activate the receptor can be tested directly. In this case, the agonists of the invention are used as controls. In addition, by using the agonist of the invention to activate the recombinant receptor, the effect of the candidate drug on this activation can be tested directly. Recombinant cells expressing the nucleic acids encoding the receptor are incubated in the assay in the presence of agonist with and without the candidate compound. A diminution in activation in the presence of the candidate will indicate an antagonist effect. Conversely, the ability of a candidate drug to reverse the antagonist effects of an antagonist of the invention may also be tested.

In an alternative to measuring calcium mobilization, the voltage clamp assay can be used as a measure for receptor activation. Agonist-induced inward chloride currents are measured in voltage-clamped oocytes expressing C140 receptor encoding cRNA or cells expressing DNA from recombinant expressions systems essentially as previously described (Julius, D., et al, Science (1988) 241:558-563) except that the single electrode voltage-clamp technique is employed.

Detection of Activated Receptors

In one embodiment, the availability of the recombinant C140 receptor protein permits production of antibodies which are immunospecific to the activated form of the receptor which can then be used for diagnostic imaging of activated receptors *in vivo*. These antibodies are produced either to the activated form of the receptor produced recombinantly, or to the peptide representing the "new amino terminal" peptide described herein. The resulting antibodies, or the immunospecific fragments thereof, such as the Fab, Fab', Fab'_2 fragments are then conjugated to labels which are detected by known methods, such as radiolabels including technetium^{99} and indium^{111} or other radioactive labels as is known in the art. When injected *in vivo*, these antibodies home to the sites of activated receptor, thus permitting localization of areas containing activated receptors.

In another embodiment, the presence of the activation peptide in body fluids or in culture media can be detected and measured. Antibodies are made to the activation

peptide as described above and can be employed in standard ELISA or RIA assays to detect excess amounts of the activation peptide in, for example, urine.

Administration of Agonists and Antagonists as Pharmaceuticals

The peptides of the invention which behave as agonists are administered in conventional formulations for systemic administration as is known in the art. Typical such formulations may be found, for example, in Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton PA, latest edition.

Preferred forms of systemic administration of peptides include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can also be used. More recently, alternative means for systemic administration of peptides have been devised which include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if properly formulated in enteric or encapsulated formulations, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels and the like.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the patient's condition, and the judgment of the attending physician. Suitable dosage ranges, however, are in the range of 0.1-100 μg/kg of subject. Wide variations in the needed dosage, however,

are to be expected in view of the variety of peptides available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization as is well understood in the art.

As shown hereinbelow, the agonists of the invention behave as antihypotensives; antagonists have the opposite effect. Thus, patients whose blood pressure needs to be raised or lowered benefit by the administration of the suitable peptide.

In addition, the agonists have anti-inflammatory and wound healing properties.

Antisense, Triple Helix and Gene Therapy Aspects

The constitutive expression of antisense RNA in cells has been shown to inhibit the expression of about 20 different genes in mammals and plants, and the list continually grows (Hambor, J.E. *et al.*, J. Exp. Med. 168:1237-1245 (1988); Holt, J.T. *et al.*, Proc. Nat. Acad. Sci. 83:4794-4798 (1986); Izant, J.G. *et al.*, Cell 36:1007-1015 (1984); Izant, J. G., *et al.*, Science 229:345-352 (1985) and De Benedetti, A. *et al.*, Proc. Nat. Acad. Sci. 84:658-662 (1987)). Possible mechanisms for the antisense effect are the blockage of translation or prevention of splicing, both of which have been observed in vitro. Interference with splicing allows the use of intron sequences (Munroe, S.H., EMBO. J. 7:2523-2532 (1988) which should be less conserved and therefore result in greater specificity in inhibiting expression of a

protein of one species but not its homologue in another species.

Therapeutic gene regulation is accomplished using the "antisense" approach, in which the function of a target gene in a cell or organism is blocked, by transfection of DNA, preferably an oligonucleotide, encoding antisense RNA which acts specifically to inhibit expression of the particular target gene. The sequence of the antisense DNA is designed to result in a full or preferably partial antisense RNA transcript which is substantially complementary to a segment of the gene or mRNA which it is intended to inhibit. The complementarity must be sufficient so that the antisense RNA can hybridize to the target gene (or mRNA) and inhibit the target gene's function, regardless of whether the action is at the level of splicing, transcription or translation. The degree of inhibition, readily discernible by one of ordinary skill in the art without undue experimentation, must be sufficient to inhibit, or render the cell incapable of expressing, the target gene. One of ordinary skill in the art will recognize that the antisense RNA approach is but one of a number of known mechanisms which can be employed to block specific gene expression.

By the term "antisense" is intended an RNA sequence, as well as a DNA sequence coding therefor, which is sufficiently complementary to a particular mRNA molecule for which the antisense RNA is specific to cause molecular hybridization between the antisense RNA and the mRNA such that translation of the mRNA is inhibited. Such hybridization must occur under in vivo conditions, that is, inside the cell. The action of the antisense RNA

results in specific inhibition of gene expression in the cell. (See: Albers, B. *et al.*, *MOLECULAR BIOLOGY OF THE CELL*, 2nd Ed., Garland Publishing, Inc., New York, NY (1989), in particular, pages 195-196.

The antisense RNA of the present invention may be hybridizable to any of several portions of a target mRNA, including the coding sequence, a 3' or 5' untranslated region, or other intronic sequences. A preferred antisense RNA is that complementary to the human C140 receptor mRNA. As is readily discernible by one of skill in the art, the minimal amount of homology required by the present invention is that sufficient to result in hybridization to the specific target mRNA and inhibition of its translation or function while not affecting function of other mRNA molecules and the expression of other genes.

Antisense RNA is delivered to a cell by transformation or transfection with a vector into which has been placed DNA encoding the antisense RNA with the appropriate regulatory sequences, including a promoter, to result in expression of the antisense RNA in a host cell.

"Triple helix" or "triplex" approaches involve production of synthetic oligonucleotides which bind to the major groove of a duplex DNA to form a colinear triplex. Such triplex formation can regulate and inhibit cellular growth. See, for example: Hogan *et al.*, U.S. Patent 5, 176,996; Cohen, J.S. *et al.*, *Sci. Amer.*, Dec. 1994, p. 76-82; Helene, C., *Anticancer Drug Design* 6:569-584 (1991); Maher III, L. J. *et al.*, *Antisense Res. Devel.* 1:227-281 (Fall 1991); Crook, S.T. *et al.* eds., *ANTISENSE RESEARCH AND APPLICATIONS*, CRC Press, 1993. It is based in part on

the discovery that a DNA oligonucleotide can bind by triplex formation to a duplex DNA target in a gene regulatory region, thereby repressing transcription initiation (Cooney M. *et. al.* (1988) *Science 241*:456). The present invention utilizes methods such as those of Hogan *et al.*, *supra* (herein incorporated by reference in its entirety), to designing oligonucleotides which will bind tightly and specifically to a duplex DNA target comprising part of the C140 receptor-encoding DNA or a regulatory sequence thereof. Such triplex oligonucleotides can therefore be used as a class of drug molecules to selectively manipulate the expression of this gene.

Thus the present invention is directed to providing to a cell or administering to a subject a synthetic oligonucleotide in sufficient quantity for cellular uptake and binding to a DNA duplex of the target C140 receptor-coding DNA sequence or a regulatory sequence thereof, such that the oligonucleotide binds to the DNA duplex to form a colinear triplex. This method is used to inhibit expression of the receptor on cells *in vitro* or *in vivo*. Preferably the target sequence is positioned within the DNA domain adjacent to the RNA transcription origin. This method can also be used to inhibit growth of cells which is dependent on expression of this receptor. The method may also be used to alter the relative amounts or proportions of the C140 receptor expressed on cells or tissues by administering such a triplex-forming synthetic oligonucleotide.

The following examples are intended to illustrate but not to limit the invention.

Example 1

Isolation of the Gene Encoding Murine C140 Receptor

A mouse cosmid genomic library (obtained from Dr. R.A. Wetsel, Washington University School of Medicine, St. Louis, Missouri and described in Wetsel, R.A. *et al.*, J Biol Chem (1990) 265:2435-2440) was screened with two ^{32}P-labeled oligonucleotides corresponding to bp 190-249 and 742-801, respectively, of the bovine substance K receptor cDNA (Masu, Y. *et al.*, Nature (1987) 329:836-838). The hybridization conditions are 5 x SSC, 5 x Denhardt's, 0.1% SDS, 0.1 mg/ml sperm DNA, 10^6 cpm/ml of labeled oligonucleotides, 60°C overnight, followed by washing with 1 x SSC, 0.1% SDS at 60°C.

In one of the clones isolated (C140) the hybridizing region was localized to a 3.7 kb PstI fragment. This fragment was subcloned into the commercially available pBluescript vector. The hybridizing and adjacent regions were sequenced in both orientations by the Sanger chain termination method. Figure 1A-1B shows both the nucleotide sequence and the deduced amino acid sequence of the mouse C140 receptor. The tentative signal sequence (SP) and the seven transmembrane regions are overlined, potential asparagine-linked glycosylation sites are marked with bold arrows, and the putative protease receptor cleavage site at Arg34-Ser35 is marked with an open arrow.

Example 2

Isolation of the Gene Encoding Human C140 Receptor

The availability of genomic DNA encoding the mouse protease C140 receptor permitted the retrieval of the corresponding human gene. A human genomic library cloned in the vector EMBL3 was screened at exactly the conditions in Example 1 using the entire coding region of the murine clone as a probe. The recovered human gene including the DNA sequence and the deduced amino acid sequence are shown in Figure 2A-2B. Subsequent experiments indicated that the human C140 gene is located in the same region of the long arm of chromosome number 5 (5q12-5q13) as has been reported for the human thrombin receptor gene.

In addition, a 1.1 kb genomic DNA fragment was obtained from Genome Systems Inc., commercial screening service as was PCR-positive with a primer pair that generates a fragment spanning 350-nucleotides of the human C140 protein coding region. A 1.1 kb bamH1 fragment was subcloned and sequenced and found to contain 800-nucleotides of promoter sequence. The promoter lacks both a TATA box and a CAAT box but is rich in G's and C's; features common to promoters of many housekeeping genes. Two binding elements specific for SP1 and AP2 were identified.

Example 3

Comparison of Related G-Protein Receptors

As shown in Figure 3, the deduced amino acid sequence of the human protease C140 receptor shows extensive similarity (>90%) to the mouse sequence.

Figure 5 shows an amino acid sequence alignment between the mouse C140 receptor and the related G-protein receptor human thrombin receptor (Coughlin, S. Cell). The tentative signal sequences (SP), transmembrane regions, and protease cleavage sites are marked.

Example 4

Recovery of Mouse C140 cDNA

A cDNA library from a mouse stomach was constructed in λ gt10 and screened with a probe encompassing the C1040 genomic DNA. A single phage clone was isolated and cut with EcoRI. The insert was cloned into pBluescript and pSG5 and sequenced.

The isolated cDNA was 2732 nucleotides long including a 16 base polyA-stretch; 5′ RACE resulted in the addition of only 27 bases to the 5′ end. The 5′ end of the apparent coding region differs from the 5′ end of the open reading frame of genomic DNA; it is believed that the 5′ end of the cDNA is correct. The complete nucleotide sequence and deduced amino acid sequence of murine cDNA encoding C140 is shown in Figure 10A-10B.

Example 5

Recovery of Human cDNA Encoding C140

A human intestinal tumor cDNA library was subjected to PCR using primers designed from the genomic clone of Example 2 and the amplified fragment was cloned in pSG5 and sequenced. The nucleotide sequence and deduced amino acid sequence are shown in Figure 11A-11B. There are four amino acid differences between the cDNA encoded sequence

and that encoded by the genomic DNA as is shown in Figure 11A-11B.

Example 6

Activation of Protease C140 Receptor in Oocytes

Both native and mutant C140 receptors were produced in oocytes and activated with a peptide mimicking the new amino-terminus", or by the proteolytic enzyme trypsin (which cleaves the extracellular region). Native receptors were produced by cloning the coding region of the receptor gene, using the polymerase chain reaction, into the expression vector pSG-5 (Green, S. *et al.*, Nucleic Acid Res (1988) 16:369). The orientation and integrity of the cloned coding region was verified by determining the nucleotide sequence with the Sanger chain-termination method. Site-directed mutagenesis was employed to construct mutant receptors in the pSG-5. Three mutant receptors were made, in which serine-35 was replaced with proline, arginine, and histidine, respectively. The nucleotide sequences of the three mutants was verified as above.

In order to produce the receptor at the surface of oocytes, cRNA encoding the receptor was produced as follows. pSG-5 C140 plasmid DNA was made linear by digestion with XbaI, and capped cRNA was produced *in vitro* using T7 RNA polymerase (Krieg and Melton, Meth Enzymol (1987) 155:397-415, which reference is hereby incorporateds by reference in its entirety).

Oocytes from Xenopus laevis were harvested and prepared using published techniques (Coleman, A., in Hames, B.D., and Higgins, S.J., eds, Transcription and

Translation: A Practical Approach, IRL Press, pp. 271-302; Williams, J.A., *et al.* Proc Natl Acad Sci USA (1988) 85:4939-4943]. To remove follicular cells, oocytes were incubated for 1.5 h with shaking in calcium-free Barth's containing 2 mg/ml each of collagenase 1A and hyaluronidase 1S. The oocytes were then washed five times in regular Barth's and incubated at 18°C in Barth's medium containing 100 U/ml penicillin, 100μg/ml streptomycin, and 2.5 mM sodium pyruvate. Stage V oocytes were selected and injected with 30 nl of cRNA (0.33 μg/μl water) or water alone, and then incubated with 0.25 ml of medium in groups of four/well in a 96-well culture plate. After 36 hours the oocytes were incubated with ^{45}Ca (250 μCi/ml). After 12 h incubation the oocytes were washed and 0.2 ml of medium added and replaced every five minutes. The harvested medium was analyzed by scintillation counting. After five replacements to determine the baseline release of ^{45}Ca, test medium with the agonist, e.g. SLIGRL, was added and the evoked ^{45}Ca-release determined.

Oocytes were injected with capped cRNA (ca 10 ng) encoding wild-type mouse C140 receptor (WT) or either of the three mutant receptors 35Pro, 35Arg and 35His. After 36 hours, cRNA-injected and control water-injected, oocytes were loaded with ^{45}Ca, and 12 hours thereafter peptide or trypsin-induced ^{45}Ca release were determined as described above. The peptide SLIGRL was added at 100 μM, and trypsin at 300 pM. The stimulation with the peptide was done on the same group of oocytes after the stimulation with trypsin. The data shown in Table 1 represent the mean of three replicate determinations, and

denotes the increase compared to oocytes injected with water.

Table 1

Receptor	Agonist	Fold increase in ^{45}Ca
WT	Trypsin	6.6
35Pro	Trypsin	0
35Arg	Trypsin	0
35His	Trypsin	0
WT	SLIGRL	11
35Pro	SLIGRL	23
35Arg	SLIGRL	15
35His	SLIGRL	23

As shown in Table 1, the agonist peptide SLIGRL was able to activate both the wild-type and mutated receptors. On the other hand, trypsin, which can activate only by cleavage of the extracellular domain, is able only to activate the wild-type receptor.

Example 7

Activation of the C140 Receptor by Different Agonist Peptides

Various peptides were tested at 100μM in the assay above using wild-type mouse C140 receptor, expressed in oocytes. The results are shown in Table 2.

Table 2

Peptide	Fold Increase in ^{45}Ca
SLIGRL	15
SLIGRA	8.5
SLIGAL	0
SLIARL	4.3
SLAGRL	0
SAIGRL	0
ALIGRL	1.3
SFFLRW	1.7

The "native" peptide SLIGRL is most effective; replacing L at position 6 with alanine lowers but does not destroy activity. Positions 2 and 3 are more sensitive. Position 1 tolerates substitution with alanine but decreases the activity by a factor of 10; the activity of this agonist is comparable to the analogous thrombin receptor agonist SFFLRW.

Example 8

Expression of C140 Receptor in Various Tissues

Poly(A)+RNA was prepared from mouse tissues, resolved on a 1.2% agarose gel containing 50% formamide and blotted onto Hybond C extra membrane (Amersham). The blot was hybridized with a ^{32}P-labeled "random priming probe" directed against the whole coding region of murine C140 receptor. The probe was hybridized at 42°C for 48 hr then successively washed at 20°C in 1 X SSC, 0.1% SDS twice, 5 min each time, then at 65°C in 1 X SSC, again twice for 20 min each time, and then 0.1 X SSC, 0.1% SDS twice for 20 min each time. The resulting membrane was

autoradiographed for 5 days at -80°C with an intensifying screen.

The results, shown in Figure 6 indicate that kidney and small intestine, but not spleen, contain mRNA encoding C140. In Figure 6, where each lane contains 10 µg RNA, lane A is derived from spleen, lane B from kidney and lane C from small intestine.

Example 9

Expression of C140 Transcripts In Mice

In situ hybridization using ^{35}S RNA probes was used to localize C140 transcripts in mouse embryogenesis and in adult mouse tissues. A strong signal was found in the gastrointestinal tract at 11.5 days; at 14 days there was strong hybridization to epithelial structures in the nasopharynx, stomach-intestine, skin and endothelial cells in larger vessels. There was some hybridization in the liver and sclerotoma but no signal in muscle or CNS. At 17 days, the signals in the sclerotoma had disappeared and additional epithelial structures showed hybridization including the esophagus, kidney glomeruli, lung, hair follicles and epidermis.

In newborns, the signals found at 17 days were retained and additional signals were found in the thymic medulla and kidney medulla. Adults showed transcripts in the mucosa of stomach, intestine and colon, white pulp of the spleen, thymus and kidney medulla. Again, there were no signals in the CNS, liver, lung or adrenal gland. Figure 12 shows the results of *in situ* hybridization in a sectioned newborn mouse using these probes.

Example 10

Expression of C140 Transcripts In Human Tissues

Figure 13 shows the results of a Northern blot of total RNA from human cell lines hybridized to a human C140 receptor probe. Ten mg of total RNA was used. Hybridization was obtained in RNA from stomach (lane 1), Ca-Co-2 cells (lane 2); HT-29 cells (lane 3), A498 cells (lane 5), 5637 cells (lane 8); skin keratinocytes (lane 12), and HUVEC (lanes 13 and 14). No hybridization was detected in HuTu80 cells, J82 cells, MCF-7, HeLa or NCI 12 cells (lanes 4, 6, 9 and 10).

Example 11

Determination of Hypotensive Activity of C140 Agonists

The C140 agonist SLIGRL was injected in 0.2 ml buffer at various concentrations into rat femoral vein and the arterial pressure was monitored. The results of various concentrations are shown in Figure 7.

The trace in Figure 7 shows that even at 0.1 mM an appreciable decrease in blood pressure occurred; larger decreases were observed at 1 mM concentration.

This effect was also shown by observing vasodilation as a result of stimulation of the rat femoral vein with the above agonist. Adult Sprague-Dawley rats were killed by exsanguination during diethylether anesthesia and the femoral vein was removed and dissected free from fat and connective tissue. Circular preparations of the vein were mounted in an organ bath (5 ml) on two L-formed metal holders (0.2 mm diameter). One of the metal holders was screwed into one of the levers of

a Grass FTO C force displacement transducer. The bathing liquid was Kreb's Ringer solution containing 118 mM NaCl, 4.7 mM KCl, 2.5 mM $CaCl_2$, 1.2 mM $MgSO_4$, 24.8 mM $NaHCO_3$, 1.2 mM KH_2PO_4 and 5.6 mM glucose. The bathing fluid was continuously treated with 88.5% oxygen-11.5% CO_2; the temperature was held at 37°C. The endothelium was removed by bubbling CO_2 through the vessels. The basal tension was between 7.5 and 12 mN. The preparations were equilibrated for at least 1 hr before application of agonist and control substances.

The results of these determinations are shown in Figure 8a and 8b. As shown in Figure 8a, contraction induced by application of $PGF_{2\alpha}$ at 3 X 10^{-5} M is relaxed by administration of 10^{-5} M agonist. The results in Figure 8a were obtained using the vein with the endothelium still present.

In Figure 8b, the endothelium has been removed. In an analogous experiment, the contraction induced by 3 X 10^{-5} M $PGF_{2\alpha}$ is not counteracted by 10^{-5} M agonist or by 10^{-5} M acetylcholine.

Example 8

Activation of Recombinant C140 Receptor by Plasmin and Kallikrein

Figures 9a and 9b show the ability of plasmin and kallikrein respectively to activate oocytes injected with C140 cRNA (open circles) or water (crosses) as control. Figure 9c shows the ability of trypsin to activate frog oocytes injected with C140 receptor cRNA (filled circles) or substance K receptor cRNA (open

circles). Trypsin clearly has a differential effect on the C140 receptor-injected oocytes.

All references cited and mentioned above, including patents, journal articles and texts, are all incorporated by reference herein, whether expressly incorporated or not.

Having now fully described this invention, it will be appreciated by those skilled in the art that the same can be performed within a wide range of equivalent parameters, concentrations, and conditions without departing from the spirit and scope of the invention and without undue experimentation.

While this invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications. This application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth as follows in the scope of the appended claims.